Original Communication

Increasing the confidence in half-sibship determination based upon 15 STR loci

Chang En Pu MS (Scientist Specialist and Section Chief) [a,*],
Adrian Linacre PhD (Senior Lecturer) [b]

[a] *Scientific and Technical Research Center, Ministry Justice Investigation Bureau, Taipei County 231, Taiwan, ROC*
[b] *Centre for Forensic Science, WestCHEM, University of Strathclyde, Glasgow, UK*

Abstract

A half-sibship relationship is when two siblings share only one parent. It may be necessary to determine if two individuals are half-siblings in cases of immigration, inheritance, genetic counseling or the identification of human remains. In such instances a combined half-sibship index (CHSI) can be calculated. Support for this kinship is also based upon the number of shared-alleles at DNA loci. We report on the combination of the calculation of CHSI with the all-shared-alleles (ASA) to enhance the specificity of any half-sibship test. The 15 STR loci (including CODIS 13) that comprise the Identifiler® loci were applied to three populations using 355,620 simulated pairs of half-siblings and 178,815 unrelated pairs. Based upon the data obtained, the sensitivity and specificity can be evaluated to determine the existence of half-sibship. This report highlights the uncertainty problems inherent in this form of indirect kinship testing and recommends a combination evaluation of CHSI and ASA.
© 2008 Elsevier Ltd and FFLM. All rights reserved.

Keywords: Short tandem repeat (STR); Kinship; Half-sibship; Combined half-sibship index (CHSI); Sensitivity; Specificity

1. Introduction

A claim of half-sibling relationship may be made in cases of inheritance, immigration[1] and genetic counseling.[2] A genetic link as a half-sibling may assist in the identification of human remains if no closer genetic relative is available for testing. A half-sibling is defined as two individuals sharing one parent, this may be a mother when it is termed uterine, or one father when it is called agnate or consanguine. Theoretically there is a 0.5 chance that two half-siblings will share 1 allele at any one locus and a 0.5 chance that they will share no alleles.[4] The confidence that two individuals are half-siblings, or not, will increase if more loci are examined, or if there are more potential half-siblings for comparison. When determining the probability whether or not two individuals are half-siblings it is necessary to consider the allele frequencies of any alleles that are shared. From this calculation a combined half-sibship indices (CHSI)[3] is reported. If CHSI is less than 1 then the evidence supports two individuals as not being related as half-siblings, otherwise the data supports the existence of a half-sibling relationship. There remains much uncertainty after this calculation.[1,3,4] Uncertainty remains with the evaluation of the resulting LR figure in terms of its reliability in determining whether two samples have come from half-siblings.

We report on the use 15 STR, that comprise the Identifiler® loci (including the 13 CODIS loci), to study three populations containing 355,620 simulated half-sibling pairs and 178,815 random pairs generated from DNA profiles created within the populations. Using this large number of sample pairs, it is possible to evaluate the sensitivity and specificity for the 15 STR core set in discriminating between half-siblings and random pairs. The discriminating power combining the CHSI and all-shared-alleles (ASA) is reported that would assist a laboratory with an interpretation of test results.

* Corresponding author. Tel.: +886 2 29112241x3740; fax: +886 2 29138599.
E-mail address: pu_macros@yahoo.com.tw (C.E. Pu).

2. Materials and methods

450 STR DNA profiles from random Chinese individuals within the Taiwan population were obtained using the ABI AmpF*l*STR Identifiler® PCR Amplification Kit (Applied Biosystem, Foster City, CA, USA).

The profiles were processed by Microsoft Excel Macros controlled by built in Visual Basic a program written by authors of this study. Every individual of the population was paired with every other individual to form random pairs, e.g. for Taiwan population in this study $(450 \times 449)/2$, equaling 101,025 pairs, were made. For generating half-siblings, every individual in the population was paired with two other individuals, e.g., $(449 \times 448)/2$, equaling 100,576 triples (one person with two mates). One offspring was generated from each of the two pairs within a triple, resulting in 201,152 half-sibling pairs. DNA profiles from the 15 STR loci from a Caucasian ($n = 301$) and American African ($n = 256$) population were obtained from Short Tandem Repeat DNA Internet Database.[6] These data were processed in the same way as those for the Taiwanese population to generate half-sibling pairs and random pairs. For the Caucasian population 89,700 half-sibling pairs were generated and 45,150 random pairs were obtained and for African American population 64,770 half-sibling pairs and 32,640 random pairs were generated. The CHSI were calculated according to standard formulae,[3] and allele frequency tables used for Caucasian and African American populations were downloaded from the same origin as the DNA profiles[6] and the frequency table for the Chinese population was from previous studies.[7–9]

All the frequencies of alleles were adjusted by using 5/2N rule[10] for this study.

The rate of false negatives was calculated as the percentage of real half-sibship that could be excluded based upon any given cutoff point of CHSI. The rate of false positives equaled the percentage of random pairs of DNA profiles where their CHSI was greater than any chosen CHSI threshold value. The sensitivity of the test is based upon 1 – the % of false negatives and the specificity is based upon 1 – the % of false positives.[11]

ASA were determined by counting all the alleles shared by paired profiles. In the case of both alleles at any one locus being shared a score of 2 is registered. If 1 allele is in common a score of 1 is registered.

The above formulae were also applied on four situations, like low CHSI with low ASA, low CHSI with high ASA, high CHSI with low ASA and high CHSI with high ASA for real half-siblings and random pairs to evaluate the synergy effects combing both CHSI and ASA.

3. Results and discussion

3.1. Ratio distribution of CHSI for three populations by using 15 STR core set

The CHSI ratio distribution of simulated half-sibling pairs and random pairs for three populations is shown in Fig. 1. Bipolar models with a widespread ratio distribution were found for all of the three populations. Simulated siblings pairs with very low CHSI values (1.40E−03 for the Chinese population) and random pairs with very high
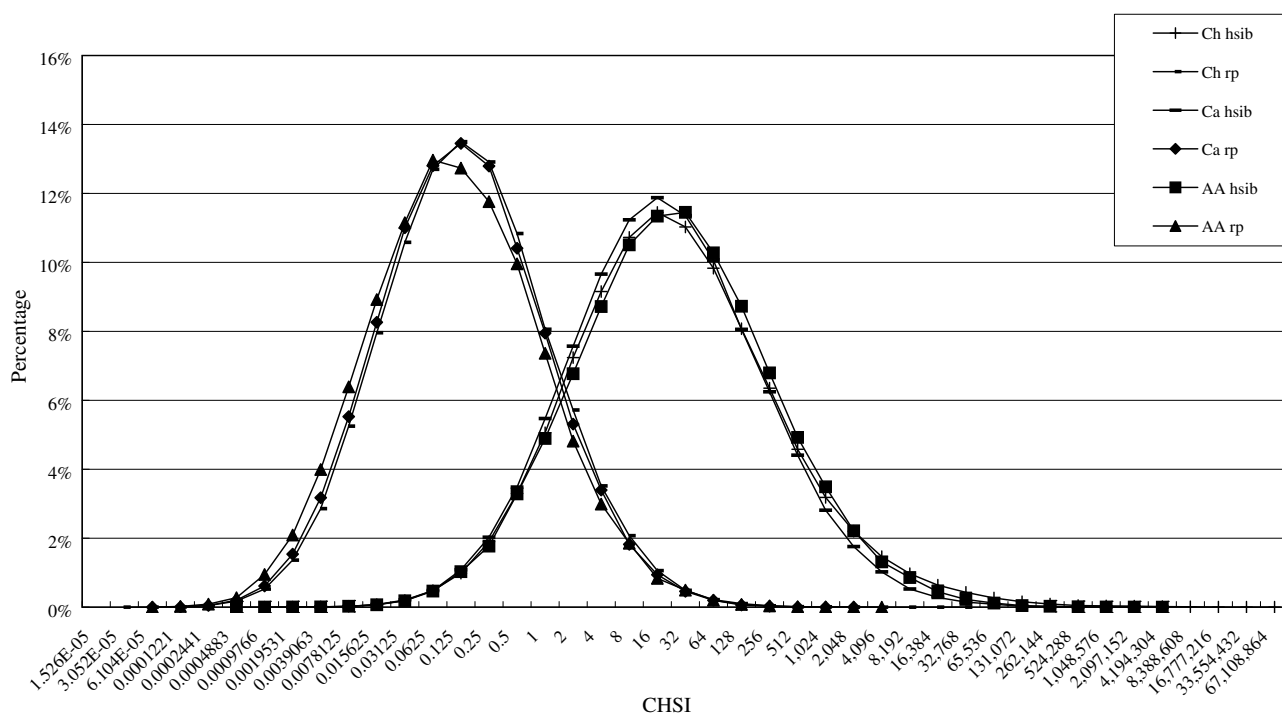

Fig. 1. CHSI ratio distribution for three populations using the 15 STR system (Ch: Chinese, Ca: Caucasians, AA: African Americans, hsib: half-siblings; rp: random pairs).

CHSI values (2.29E+05 for African American population) were observed in this study, together with other extreme CHSI values are reported in Table 1. From the data obtained for the 15 STR loci (Table 2), 12.02%, 12.86% and 11.73% of CHSI values from simulated half-sibling pairs were found to be less than 1 for Chinese, Caucasians and African Americans populations, respectively. Random pairs that produced CHSI values greater than 1 for these populations were 13.24%, 12.22% and 11.33% for these three populations.

3.2. Sensitivity and specificity under different CHSI cut-offs for 15 STR systems

The ability of the DNA test to correctly classify kinship testing results into two categories is assessed by their specificity and sensitivity.[11] In Table 2, the sensitivity and specificity using a range of CHSI values from 0.03125 to 1000 are illustrated. The CHSI requirements as cutoff values followed previous recommendations.[5] As the CHSI cutoff values increased there was a corresponding decrease in sensitivity and increase in specificity. According to Table 2, when adopting CHSI value of 1 as the cutoff value, then in 15 STR core set system for the three populations the sensitivity was 87.98% and the specificity was 86.76% for the Chinese population, for Caucasian population the sensitiv-

Table 1
Highest and lowest CHSIs for three populations using 15 STR systems

		Pairs	Highest	Lowest
Chinese	Simulated half-siblings	201,152	1.90E+07	1.40E−03
	random pairs	101,025	1.04E+04	5.34E−05
Caucasians	Simulated half-siblings	89,700	1.35E+06	2.30E−03
	random pairs	45,150	6.90E+02	7.77E−05
African Americans	Simulated half-siblings	64,770	1.37E+06	8.13E−04
	random pairs	32,640	2.29E+05	7.15E−05

Table 2
Evaluation of sensitivity, specificity and CHSI threshold value for half-sibship determination using 15 STR system

Population	Chinese		Caucasians		African Americans	
CHSI threshold value	[a]SEN (%)	[b]SPE (%)	[a]SEN (%)	[b]SPE (%)	[a]SEN (%)	[b]SPE (%)
0.03125	99.70	28.76	99.69	30.37	99.71	33.88
0.0625	99.22	41.46	99.21	43.18	99.24	46.85
0.125	98.23	54.96	98.12	56.63	98.21	59.59
0.25	96.34	67.87	96.10	69.42	96.44	71.35
0.5	93.04	78.70	92.62	79.83	93.16	81.31
1	87.98	86.76	87.14	87.78	88.27	88.67
2	80.74	92.48	79.58	93.09	81.50	93.48
3	75.63	94.80	74.20	95.30	76.62	95.39
10	57.19	98.49	54.92	98.67	58.83	98.69
33	37.93	99.63	34.97	99.71	39.02	99.68
100	23.14	99.92	19.99	99.96	23.37	99.91
150	18.83	99.95	15.70	99.98	18.74	99.94
200	16.19	99.97	13.05	99.99	15.86	99.96
300	12.95	99.98	9.90	100.00	12.33	99.98
330	12.25	99.98	9.23	100.00	11.59	99.98
7500	9.68	99.99	6.79	100.00	8.89	99.99
1000	6.46	100.00	3.95	100.00	5.37	99.99

[a] Sensitivity: % true half-sibs with CHSI values greater than threshold.
[b] Specificity: % random pairs with CHSI values less than threshold.
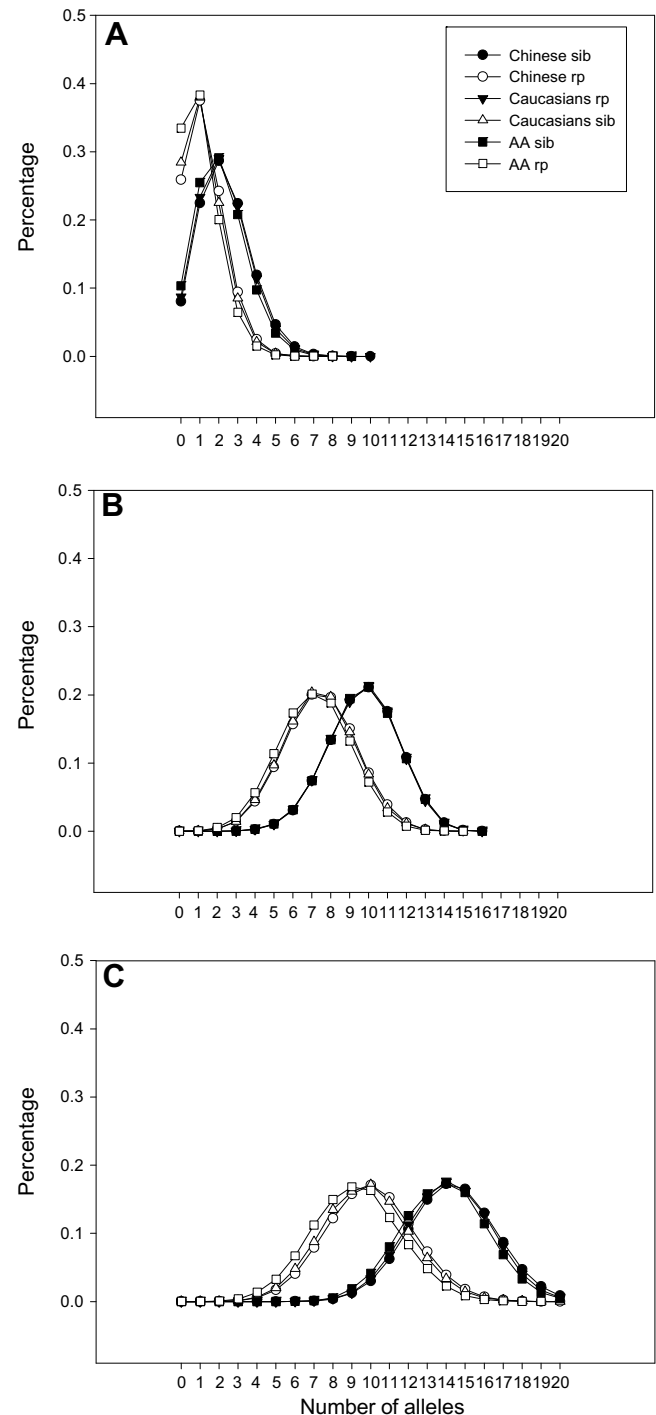

Fig. 2. Ratio distribution of allele sharing instances. A for the number of 2-allele sharing loci, B for the number of loci with one allele shared and C when all shared alleles are counted. Half siblings are denoted by hsib and random pairs by rp.

Table 3
Synergy effects of All-Shared-Alleles cutoffs and CHSI cutoffs for half-sibship determination by using 15 STR

Chinese	0.125		0.25		0.5		1		1.25		1.5		2		3		10	
ASA	SEN (%)	SPE (%)	SEN (%)	SPE (%)	SEN (%)	SPE (%)	SEN (%)	SPE (%)	SEN (%)	SPE (%)	SEN (%)	SPE (%)	SEN (%)	SPE (%)	SEN (%)	SPE (%)	SEN (%)	SPE (%)
7	98.23	54.96	96.34	67.87	93.04	78.71	87.98	86.76	85.87	88.89	84.02	90.39	80.74	92.48	75.63	94.80	57.19	98.49
8	98.22	55.03	96.33	67.89	93.04	78.71	87.98	86.76	85.87	88.89	84.02	90.39	80.74	92.48	75.63	94.80	57.19	98.49
9	98.11	55.57	96.27	68.04	93.00	78.75	87.96	86.78	85.85	88.90	84.00	90.40	80.73	92.49	75.62	94.81	57.19	98.49
10	97.42	58.16	95.82	68.98	92.74	79.07	87.80	86.87	85.72	88.98	83.89	90.46	80.64	92.52	75.56	94.82	57.17	98.49
11	94.96	65.15	93.92	72.24	91.45	80.27	86.98	87.30	85.02	89.28	83.28	90.68	80.13	92.66	75.18	94.89	57.03	98.50
12	89.00	75.86	88.57	78.72	87.15	83.38	83.86	88.58	82.26	90.19	80.79	91.39	78.07	93.14	73.62	95.16	56.39	98.54
13	78.46	86.14	78.37	86.78	77.84	88.49	76.13	91.22	75.15	92.23	74.19	93.03	72.29	94.23	68.93	95.75	54.22	98.64
14	63.57	93.34	63.56	93.41	63.46	93.73	62.91	94.60	62.53	95.00	62.11	95.35	61.15	95.96	59.21	96.84	48.89	98.84
15	46.36	97.24	46.36	97.24	46.35	97.29	46.24	97.43	46.14	97.52	46.01	97.62	45.71	97.81	44.96	98.16	39.51	99.17
16	29.86	99.02	29.86	99.02	29.86	99.02	29.85	99.03	29.84	99.04	29.82	99.06	29.76	99.10	29.58	99.17	27.53	99.52
17	16.86	99.70	16.86	99.70	16.86	99.70	16.86	99.71	16.86	99.71	16.86	99.71	16.85	99.71	16.83	99.72	16.29	99.81
18	8.18	99.93	8.18	99.93	8.18	99.93	8.18	99.93	8.18	99.93	8.18	99.93	8.18	99.93	8.18	99.94	8.10	99.94
Caucasian																		
7	97.47	56.63	95.46	69.42	92.00	79.83	86.57	87.78	84.34	89.69	82.46	91.15	79.05	93.09	73.71	95.30	54.55	98.67
8	97.46	56.70	95.45	69.44	92.00	79.84	86.56	87.78	84.34	89.69	82.46	91.15	79.05	93.09	73.71	95.30	54.55	98.67
9	97.31	57.26	95.37	69.57	91.95	79.85	86.54	87.78	84.32	89.69	82.44	91.15	79.03	93.09	73.70	95.30	54.55	98.67
10	96.57	60.16	94.91	70.55	91.69	80.12	86.40	87.85	84.21	89.73	82.35	91.17	78.96	93.10	73.66	95.30	54.54	98.67
11	93.78	67.91	92.75	74.22	90.21	81.54	85.52	88.33	83.46	90.07	81.71	91.40	78.45	93.25	73.30	95.37	54.44	98.68
12	87.29	78.64	86.88	81.01	85.46	84.94	82.15	89.70	80.52	91.06	79.08	92.14	76.29	93.70	71.67	95.60	53.85	98.71
13	76.17	88.10	76.09	88.61	75.59	89.96	73.98	92.27	73.03	93.09	72.12	93.80	70.22	94.86	66.80	96.25	51.75	98.83
14	60.90	94.41	60.89	94.46	60.81	94.73	60.31	95.44	59.94	95.77	59.53	96.02	58.64	96.54	56.73	97.27	46.35	99.01
15	43.31	97.76	43.31	97.76	43.30	97.79	43.22	97.90	43.15	97.96	43.05	98.01	42.78	98.15	42.09	98.42	36.85	99.29
16	27.20	99.21	27.20	99.21	27.19	99.21	27.18	99.22	27.17	99.23	27.16	99.24	27.11	99.26	26.95	99.34	25.16	99.63
17	14.52	99.75	14.52	99.75	14.52	99.75	14.52	99.76	14.52	99.76	14.52	99.76	14.51	99.76	14.50	99.78	14.11	99.84
18	6.49	99.94	6.49	99.94	6.49	99.94	6.49	99.94	6.49	99.94	6.49	99.94	6.49	99.94	6.49	99.94	6.43	99.95
African Americans																		
7	97.35	59.60	95.53	71.35	92.10	81.31	87.05	88.67	85.05	90.66	83.26	91.75	80.08	93.48	75.12	95.39	57.27	98.69
8	97.32	59.74	95.52	71.38	92.09	81.31	87.05	88.67	85.04	90.66	83.26	91.75	80.08	93.48	75.12	95.39	57.27	98.69
9	97.13	60.71	95.40	71.66	92.02	81.39	87.02	88.68	85.02	90.67	83.24	91.76	80.07	93.49	75.11	95.39	57.27	98.69
10	95.82	64.72	94.43	73.27	91.42	81.90	86.66	88.81	84.73	90.74	82.97	91.80	79.86	93.52	74.97	95.40	57.23	98.69
11	92.09	73.69	91.33	77.86	89.15	83.76	85.14	89.52	83.42	91.16	81.80	92.11	78.92	93.69	74.27	95.49	56.98	98.69
12	84.38	83.73	84.14	85.04	83.05	87.63	80.56	91.20	79.28	92.41	78.03	93.13	75.73	94.36	71.79	95.89	55.96	98.73
13	71.83	91.69	71.79	91.89	71.48	92.54	70.45	93.97	69.80	94.57	69.07	94.97	67.74	95.66	65.13	96.68	52.71	98.86
14	56.13	96.48	56.13	96.49	56.07	96.55	55.78	96.86	55.56	97.06	55.30	97.18	54.72	97.44	53.47	97.88	45.72	99.09
15	38.95	98.72	38.95	98.72	38.95	98.72	38.90	98.74	38.84	98.76	38.79	98.78	38.63	98.84	38.26	98.99	34.76	99.45
16	23.19	99.55	23.19	99.55	23.19	99.55	23.18	99.55	23.18	99.55	23.16	99.56	23.14	99.56	23.07	99.59	22.05	99.70
17	11.95	99.87	11.95	99.87	11.95	99.87	11.95	99.87	11.95	99.87	11.95	99.87	11.95	99.87	11.94	99.87	11.73	99.88
18	5.18	99.97	5.18	99.97	5.18	99.97	5.18	99.97	5.18	99.97	5.18	99.97	5.18	99.97	5.18	99.97	5.16	99.97

ity was 87.14% and the specificity was 87.78% and for African American population the sensitivity was 88.27% and the specificity was 88.67%. Using a CHSI of 1 would falsely exclude more than 10% of real half-sibling, indicating the potential problems with this type of DNA testing.

3.3. Ratio distribution of allele sharing for three populations

The ratio distribution of allele-sharing for simulated half-sibling pairs and random pairs is shown in Fig. 2, three of the allele sharing situation were compared. Graphs C illustrate the ASA situation where there is greater separation between half-sibling pairs and random pairs (with smaller percentage of overlapping area in the bipolar model). These data indicate the potential value of ASA in the determination of half-sibship.

3.4. Evaluation of synergy effects of two criteria for the half-sibship determination

The CHSI vs. ASA ratio distribution was calculated for all the simulated half-siblings and random pairs, and the sensitivity and specificity for each CHSI and ASA combination was evaluated. Values of CHSI varying from 0.125 to 10 were analyzed against the values of 7–18 of ASA using the 15 STR system (Table 3). A combination of these two data sets increased the specificity in an inclusion of the relationship. A CHSI of 0.125 and ASA of 13 resulted in a specificity of 86.14%, whereas if only considering ASA the specificity was only 54.96% (Table 2) for the Taiwanese population. When CHSI = 1 and ASA = 13, the specificity increased to 91.22% instead of 86.76% for the Taiwanese population (see Table 2, when based on CHSI = 1 only), 92.27% instead of 87.78% for Caucasian population and 93.97% instead of 88.67% for African American population. Combining the two criterion increases the specificity for half-sibling determination especially for the cases of "low CHSI with high ASA", otherwise they are maybe denied as half-siblings by using CHSI cutoffs only.

There is alternative of resolving half-sibship cases with greater confidence by using more autosomal STR loci, using mitochondrial DNA analysis and using STR loci on either of the sex determining chromosomes.

4. Conclusion

The use of CHSI and all-shared-alleles together increases the specificity with a resulting increase in the confidence in half-sibship determination for low CHSI cases.

References

1. Allen RW, Fu J, Reid TM, Baird M. Considerations for the interpretation of STR results in cases of questioned half-sibship. *Transfusion* 2007;**47**:515–9.
2. <http://www.cbsnews.com/stories/2006/03/17/60minutes/main1414965.shtml> [accessed 1.06.07].
3. Wenk RE, Traver M, Chiafari FA. Determination of sibship in any two persons. *Transfusion* 1996;**36**:259–62.
4. Wenk RE, Chiafari FA. Distinguishing full siblings from half-siblings in limited pedigrees. *Transfusion* 2000;**40**:1148–9.
5. Gaytmenn R, Hildebrand DP, Sweet D, Pretty IA. Determination of the sensitivity and specificity of sibship calculations using AmpFISTR Profiler Plus. *Int J Legal Med* 2002;**116**:161–4.
6. Short Tandem Repeat DNA Internet Database. <http://www.cstl.nist.gov/div831/strbase/>.
7. Pu CE, Wu FC, Cheng CL, Wu KC, Chao CH, Li JM. DNA STR profiling of Chinese population in Taiwan determined by using an automated sequencer. *Forensic Sci Int* 1998;**97**:47–51.
8. Pu CE et al. *DNA profiling of Chinese population in Taiwan on the STR loci D8S1179, D16S539, D18S51 and D21S11 52th annual meeting*. Reno (NV, USA): American Academy of Forensic Sciences; 2000.
9. Lee JC, Chang YY, Su CW, Tzeng CH, Pu CE, Han TH, et al. The screening of STR and YSTR loci in Taiwanese han population. In: Proceedings of Taiwan academy of forensic sciences; 2006. p. 27–38 [in Chinese].
10. Butler JM. Kinship and parentage testing. In: Butler JM, editor. *Forensic DNA typing*. MA, USA: Elsevier Academic Press; 2005. p. 477.
11. Kirkwood BR, Sterne JAC. Measurement error: assessment and implications. In: Kirkwood BR, Sterne JAC, editors. *Medical statistics*. MA, USA: Blackwell Publishing; 2003. p. 429–46.